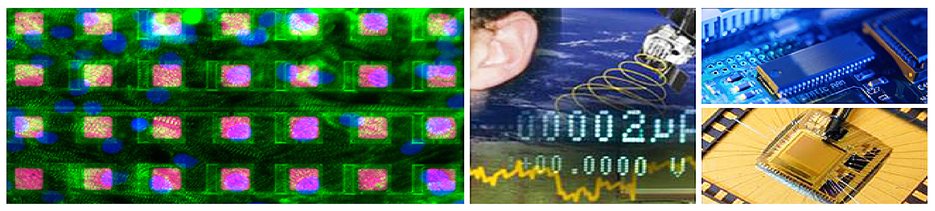
Electronics & Signal Processing Laboratory
ESPLAB
Breguet 2, CH - 2000 Neuchâtel
www.epfl.ch/esplab



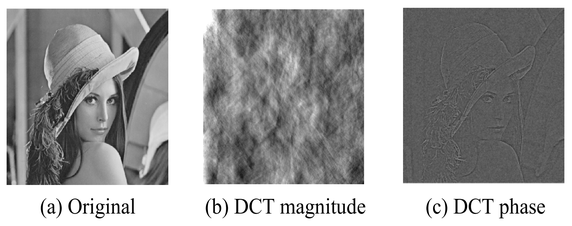
RETRIEVAL OF OCCLUDED IMAGES USING DCT PHASE AND REGION MERGING

P. Ayyalasomayajula, S. Grassi and P.-A. Farine
contact: pradyumna.ayyalasomayajula@epfl.ch

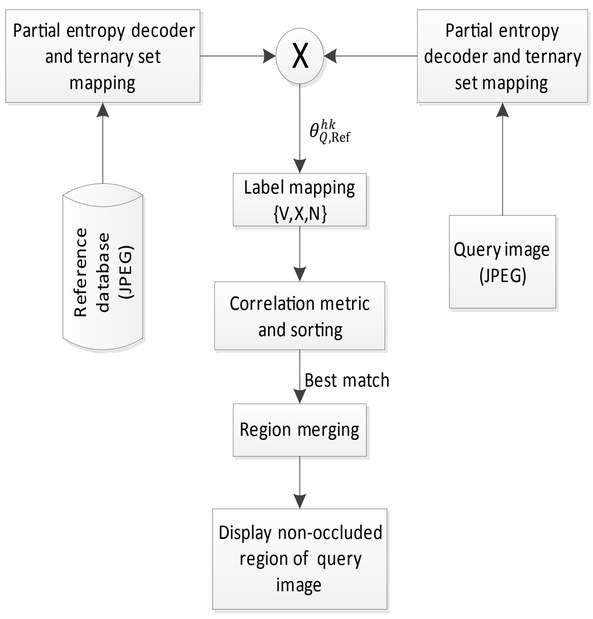
Abstract

We present an efficient method for Content Based Image Retrieval (CBIR) of occluded images using DCT-phase. The proposed method utilizes a novel correlation metric for ternary-valued DCT-phase, as well as a region merging method to reconstruct the non-occluded regions in the retrieved image. The proposed image retrieval method showed good performance when tested with different datasets containing reference images, occluded images, fused images and images with different JPEG compression ratios. Experimental evaluation also showed that the proposed image retrieval method performs better than current state of the art DCT-phase based image retrieval methods while retrieving not only occluded images but also reference images, fused images and images with different JPEG compression ratios.

Retrieval method for occluded images



(a) Original (b) DCT magnitude (c) DCT phase



Overview of the proposed image retrieval method

DCT-phase of an image and its representation

Let *s(n,m)* be a 2-dimmensional N x M sequence. The DCT of *s(n,m)* is denoted by $S_{DCT}(j,k)$. It can be expressed in terms of its absolute value, $|S_{DCT}(j,k)|$, and its corresponding phase term, $S'_{DCT}(j,k)$ as: $S_{DCT}(j,k) = |S_{DCT}(j,k)|\ S'_{DCT}(j,k)$
We represent the DCT-phase with values from the ternary set {+1,α,-1}, where the '+1' symbol corresponds to $S'_{DCT}(j,k) = +1$ (0 phase), and the '-1' symbol corresponds to $S'_{DCT}(j,k) = -1$ ('π' phase). The symbol 'α' corresponds to zero magnitude of the DCT coefficients, $|S_{DCT}(j,k)| = 0$.

Correlation metric

We first compute the ternary-valued DCT-phase of 8x8 blocks of an image I_Q with horizontal and vertical pixel resolution of W and H respectively, obtaining a W x H matrix of ternary symbols. We express this matrix as θ^{hk}_{Query}, where h = 0,1,…,(H/8)-1 and k = 0,1,…,(W/8)-1. The indexes h and k identify the corresponding 8x8 block of the image I_Q.
To correlate a query image I_Q and a reference image I_{Ref}, we first multiply element-by-element their corresponding DCT-phase arrays, θ^{hk}_{Query} and $\theta^{hk}_{Reference}$:

$$\theta^{hk}_{(Q,Ref)}(i,j) = \theta^{hk}_{Query}(i,j) \cdot \theta^{hk}_{Reference}(i,j)$$

The possible outcomes for such a multiplication where $\theta^{hk}_{Query}(i,j)$ and $\theta^{hk}_{Query}(i,j)$ can take values {+1, α, -1}, belongs to the quinary set $\{+1,\alpha,\alpha^2,-\alpha,-1\}$. An exact match is found when the outcome belongs to $\{+1,\alpha^2\}$, which forms the principal diagonal of the matrix of possible outcomes given in Table 1(a).

$\theta^{hk}_{Query}(i,j)$ \ $\theta^{hk}_{Reference}(i,j)$	+1	α	−1
+1	+1	α	−1
α	α	α^2	−α
−1	−1	−α	+1

Table 1(a): Possible outcomes for $\theta^{hk}_{(Q,Ref)}(i,j)$

V = Valid
N = Not valid
X = Don't care

$\theta^{hk}_{Query}(i,j)$ \ $\theta^{hk}_{Reference}(i,j)$	+1	α	−1
+1	V	N	N
α	N	X	N
−1	N	N	V

Table 1(b): Lables assigned to possible outcomes

Each element of the resultant arrays is assigned one of three possible "labels", Valid (V), Not-valid (N) and a Don't care (X), which correspond to {+1}, { α,-α,-1} and $\{\alpha^2\}$ respectively, as shown in Table 1(b).The similarity between the query and the reference image is calculated using the following correlation metric:

$$d(I_Q, I_{Ref}) = \frac{\sum(V)}{\sum(V) + \sum(N)}$$

Region Merging

For the array $\theta^{hk}_{(Q,Ref)}$ of a given 8x8 block to which labels (V,N,X) have been assigned , we compute the mean of the number of labels which are either valid or don't cares. That is:

$$m^{hk} = \frac{1}{64}\sum_{i=0}^{7}\sum_{j=0}^{7} 1_{\{V,X\}}\left(\theta^{hk}_{Query}(i,j)\right)$$

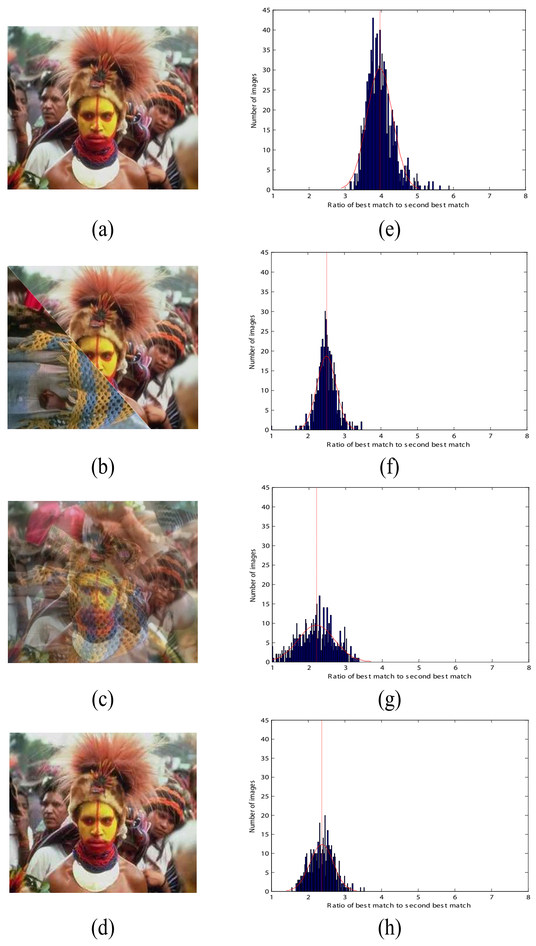
If this mean is above a given threshold, m_{thres}, the 8x8 block is classified as a matching block. After all the matching blocks are found, those which are connected to at least one adjacent matching block are retained, and the others are discharged. The non-occluded portion of the retrieved image is formed by all the retained blocks.

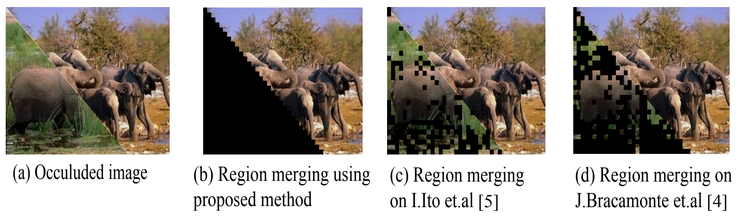
Experimental evaluation

Database : 1000 pictures from COREL photograph data set in QVGA (320x240) resolution.

The reference database was randomly divided into two datasets known as 'datasetA' and 'datasetB' each containing 500 images. A sample image from 'datasetA' is shown in (a).
An "occluded image database" was generated by occluding 37.5% of each image in 'datasetA' with part of an image in the 'datasetB': the lower-left triangular portion of the image was replaced by the same region from the occluding image. A sample image from this dataset is shown in (b).
To generate a database of fused images, called 'datasetFused', each image in the 'datasetA' was merged with an image from 'datasetB' using wavelet decomposition by taking the mean for both approximations and details coefficients (wfusimg function in Matlab). A sample image from this dataset is shown in (c).
Images in 'datasetA' were JPEG compressed with two different compression ratios (CR): 60% and 30%; creating two compressed databases 'dataset30CR' and 'dataset60CR'. A sample image from dataset30CR is shown in (d).

The ratio of the correlation of Best Match (BM) to the correlation of the Second Best Match (SBM) is used to measure the performance of the retrieval method. When testing a dataset, we calculate the mean and the standard deviation of this ratio. A high mean with low standard deviation is an indication of good performance and robustness.





(a) Occluded image (b) Region merging using proposed method (c) Region merging on I.Ito et.al [5] (d) Region merging on J.Bracamonte et.al [4]

Our region merging method successfully retrieved the non-occluded regions of the images. Figure above shows an example of our proposed region merging method (b) and of region merging using the DCT-phase representation of [5] and [4] in (c) and (d) respectively. A threshold (m_{thres}) of 0.75 was used. It is seen that the DCT-phase representation of [5] and [4] is not suitable for accurately retrieving the non-occluded portion of the image.

	Proposed method		I.Ito et.al		J.Bracamonte et.al	
Dataset	**m**	**σ**	**m**	**σ**	**m**	**σ**
Reference database	3.97	0.36	1.41	0.18	1.26	0.09
datasetOcclude	2.52	0.26	1.19	0.09	1.16	0.06
datasetFused	2.21	0.49	1.31	0.14	1.19	0.08
dataset30CR	2.37	0.32	1.04	0.03	1.13	0.05
dataset60CR	3.73	0.39	1.39	0.18	1.26	0.09

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Innovation Promotion Agency CTI